

EXHIBIT A

Structure of the λ *att* sites generated by *int*-dependent deletions

(site-specific recombination/viral integration)

RONALD H. HOESS AND ARTHUR LANDY*

Section of Microbiology and Molecular Biology, Division of Biology and Medicine, Brown University, Providence, Rhode Island 02912

Communicated by Norton D. Zinder, August 30, 1978

ABSTRACT Bacteriophage λ integrates into the chromosome of its *Escherichia coli* host by means of a site-specific recombination between a locus on the phage chromosome (phage *att* site) and a locus on the bacterial chromosome (bacterial *att* site). The nucleotide sequence of four λ *att* sites altered in site-specific recombination has been determined. The *int*-dependent deletions that generated these *att* sites have one end point within the phage *att* site and extend either to the left or to the right. As a result of the new internucleotide bond created by deletion formation, these phage have alterations in the 15-base-pair common core region. The new DNA sequences brought to the *att* sites by the deletions, designated Δ for regions to the left and Δ' for regions to the right, do not share any discernible homology with their analogous counterparts in the phage *att* site arms, P and P', respectively, or with the bacterial *att* site arms, B and B', respectively. The finding of alterations in the 15-base-pair common core region necessitates a reinterpretation of the genetic properties of these *att* sites in site-specific recombination. The structure of these sites in relation to their genetic properties can be viewed as being consistent with a model in which the only specificity elements in *int*-dependent site-specific recombination are the common core region, O, and the phage arms, P and P'.

Bacteriophage λ is an extremely useful model system for studying site-specific recombination, due in a large part to the availability of mutants directly affecting the recombination process. The mutations can occur either in genes coding for proteins required for this process (1) or in the region of the DNA (*att* sites) used as substrates for the reaction (2-5). This report describes results that begin to identify those features of the *att* site structure essential for site-specific recombination.

As first proposed by Campbell (6), upon injection into its host, *Escherichia coli*, the DNA of λ can integrate into the bacterial genome by a reciprocal recombination event. This reaction requires a functional phage gene, *int* (7-9), host factors (10 and 11), and reaction sites on the DNA of both phage and bacterium designated *attP* (POP*) and *attB* (BOB*), respectively (12). As a result of integration, λ is inserted linearly into the bacterial genome, and two prophage *att* sites, one at each end of the integrated phage, *attL* (BOP*) and *attR* (POB*), are formed. Excision, the reverse reaction, recombines *attL* with *attR* in the presence of a functional *int* gene and an additional phage gene, *xis* (13, 14), to regenerate the original *attP* and *attB* sites. Genetic evidence has shown that these two reactions occur with precision, and that the *att* sites are not altered by cycles of integration and excision (15).

Both genetic experiments (16) and DNA heteroduplex mapping analysis (17) have indicated that *attP* and *attB* are nonhomologous. Recent DNA sequence analysis has revealed, however, a 15-base-pair region of homology shared among *attP*, *attB*, *attL*, and *attR*, designated the common core, O (18). It is in this region that the crossover during integration and excision is presumed to take place. We would like to know which

features enable the *att* sites to participate in site-specific recombination. In an attempt to answer this question, we have analyzed four phages that contain different structural alterations in their *att* sites but are still proficient in certain aspects of site-specific recombination. These phages all have deletions with one end point located within the phage *att* site (17) and whose generation depended upon the presence of a functional *int* gene (3).

MATERIAL AND METHODS

Phage, λ b511*cIam*, λ b508*cIam*, and λ b522*cIam* were obtained from J. S. Parkinson (3). λ b2*cI857* was provided by F. Blattner. As a source of wild-type DNA, λ cI857*S7* was used. Phage were grown and their DNA was extracted by using described procedures (19).

Enzymes. Restriction endonucleases *EcoRI*, *HindIII*, *Mbo* II, and *Bam*HI were isolated by described procedures (20). *Hpa* II, *Hha* I, *Hae* III, *Hinf*I, and *Alu* I were purchased from New England Biolabs. Assay conditions for these enzymes have been described (21). Bacterial alkaline phosphatase was purchased from Worthington. Phosphorylation of 5' ends of DNA fragments was carried out by using polynucleotide kinase from P-L Biochemicals or from Boehringer Mannheim. DNA polymerase I (Boehringer Mannheim) in conjunction with DNase (Worthington) was used for nick translation according to the procedure of Rigby *et al.* (22).

Gel Fractionation. Restriction fragments were fractionated by electrophoresis (on either 1% agarose or polyacrylamide gels), visualized, eluted, and prepared for sequencing as described (21).

DNA Sequencing. The methods of Maxam and Gilbert (23) were used for 5'-end-labeling of DNA fragments and the subsequent nucleotide sequence determination. Sequencing gels were routinely 12% polyacrylamide with 8 M urea. For the determination of sequences from 150 to >200 bases from the 5' labeled end, 0.5-mm "thin" gels as described by Sanger and Coulson (24) were used. In addition, the gel length was increased to 70 cm and the run was >2000 V.

RESULTS

Isolation of Restriction Fragments Containing the *att* Sites. Each of the four phages used in this study (λ b2, λ b511, λ b508, and λ b522) has extensive deletions of DNA. Moreover, each deletion has one end point that has been localized within or very close to the *att* site (Fig. 1) by genetic crosses (3) and DNA heteroduplex analysis (17, 25). To identify a primary restriction fragment containing the *att* site, restriction digests of the deletion phages were compared with those of wild-type phage. As expected, each deletion phage lacked one or more restriction fragments as a result of the deletion. In addition, each

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "advertisement" in accordance with 18 U. S. C. §1734 solely to indicate this fact.

Abbreviations: POP* and BOB*, reaction sites on the DNA of phage and bacterium, respectively.

* To whom reprint requests should be addressed.

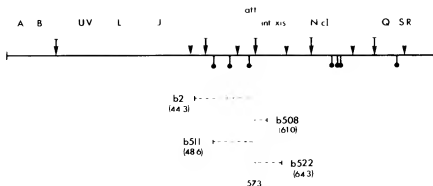


FIG. 1. Physical and genetic map of bacteriophage λ . The λ chromosome is shown as a linear molecule with various phage genes represented by letters. Below are regions of DNA deleted (---) in the four phages studied: λ b2, λ b511, λ b508, and λ b522. Deletion end points are given as % λ , assuming the cohesive ends of λ are at 0 and 100% from left to right. The end points have been fixed by electron microscope DNA heteroduplex mapping (25). Locations of *EcoRI* (▼) (26), *BamHI* (◄) (27), and *HindIII* (●) (19, 28) restriction enzyme cut sites are noted.

digest also contained a single unique fragment characteristic of that particular deletion phage. This unique fragment should contain the "novel joint" or new internucleotide bond created by the deletion. Double restriction enzyme digestions with *EcoRI* and *HindIII* were used to isolate primary fragments

from λ b508 and λ b522. *EcoRI* and *BamHI* were used to digest λ b2 and *HindIII* and *BamHI* were used for λ b511.

Primary fragments were isolated preparatively and then subjected to further restriction enzyme analysis to obtain fragments small enough for DNA sequencing. A number of

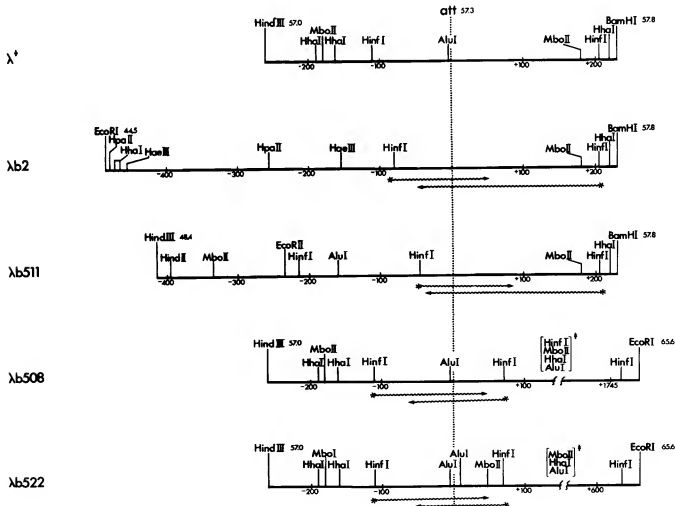


FIG. 2. Fine-structure restriction maps of *att* site containing fragments isolated from λ^+ (18) and the four deletion phages, λ b2, λ b511, λ b508, and λ b522. Map distances are given as nucleotide pairs from a fixed starting point, O, located at the center of the core (18) with positive numbers representing material to the right of *att* and negative numbers for material to the left of *att*. * indicates that relative order of these restriction sites in this region have not been determined. Labeled DNA strands used for sequence determination (---) are below each map with * representing ^{32}P -labeled 5'-termini and \rightarrow showing direction of sequence.

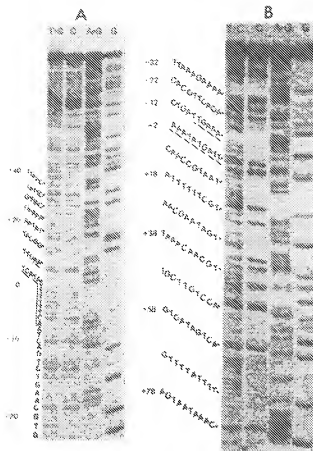


FIG. 3. Autoradiographs of two sequencing gels, showing both strands of the *HinfI*/*HinfI* fragment spanning the *att* site of λ 511. Reaction conditions for cleavages at T, C, and G were those described by Maxam and Gilbert (23). Cleavage at A was obtained by heating the sample in 1 M NaOH at 90°C for 8 min, followed by the standard treatment in 0.5 M piperidine for 30 min at 90°C. Electrophoresis was in 10% acrylamide (B) or 12% acrylamide (A). (A) Sequence derived from the 5' *HinfI* terminus located in the Δ arm of the Δ OP' phage. The core sequence has been underlined and the numbers represent distance in base pairs from the 5' terminus of the core. Base +1 then is actually 54 base pairs from the 5' terminus of this fragment. (B) Sequence derived from the complementary strand with the core sequence again underlined. In this case the base labeled +6 is 200 bases from the 5' terminus located in the P arm of the phage.

methods were used to generate the restriction fragment maps of the primary fragments shown in Fig. 2. In some cases the partial mapping procedure of Smith and Birnstiel (29) was used. As an example, the 5' *EcoRI* end of the *EcoRI*/*BamHI* primary fragment of λ b2 was labeled with 32 P. This fragment was then digested with various enzymes under conditions such that digestion does not go to completion. The number and spacing of fragments generated by partial digestion as seen by autoradiography could then be used to construct an order of fragments obtained in a complete digest. We were also aided in ordering restriction fragments by prior knowledge of the restriction map of the wild-type phage *att* site (18). Note that all of the primary fragments contain either a wild-type P or P' arm.

Nucleotide Sequence Analysis of the Hybrid *att* Sites. The arrows in Fig. 2 indicate the direction in which DNA was sequenced from the 32 P-labeled 5'-end by the procedure of Maxam and Gilbert (23). *HinfI* fragments which spanned the *att* site in all of the phages analyzed were kinase-labeled and strand separated, and both strands were sequenced. As an example, the two gels in Fig. 3 show complementary portions of the *HinfI* fragment spanning the *att* site in λ 511.

It is worth noting that, although the phages we have used for sequencing are of independent origin from those sequenced by Landy and Ross (18), our sequence data (Fig. 4) for the P and P' arms are in complete agreement. We do not find the discrepancies noted by Davies *et al.* (30) at positions +18, +35, and +63 (Fig. 3A) in their sequence analysis of the P' arm. In addition, our sequencing data have provided a complement for a short stretch of 17 bases at positions +35 to +52 in the P' arm (18).

DISCUSSION

Although the general model proposed by Campbell (6) for site-specific recombination in λ has been shown to be correct, the detailed mechanism for this process at the molecular level is not yet understood. In particular, little is known about the specificity encoded by the DNA substrates for this reaction. We have reported here the analysis of four mutant *att* sites whose recombination frequencies in site-specific recombination differ from the frequency of a wild-type phage *att* site, POP'. Whereas POP' recombines efficiently with BOB', neither the *att* sites of the leftward deletion phages (Δ OP') nor the rightward deletion phages (POA') recombine efficiently with BOB' (3, 31). Despite this difference, these mutants do not represent a null phenotype with respect to site-specific recombination, because each is capable of recombining efficiently with certain *att* sites other than BOB' (3, 31, 32). In fact, it has been possible to characterize the mutants on the basis of their relative efficiencies in *int*-dependent recombination with several different *att* sites. Because of this, the prevailing hypothesis has been that the mutants, like the wild-type *att* sites, contained a common core region, O, and that the measurable differences were due to different flanking sequences Δ and Δ' .

From our analysis we find that indeed there is an identifiable common core sequence in Δ OP' and POA' *att* sites, but it is not identical to that of wild-type (Fig. 5). Thus, it is not possible at this time to attribute the characteristic genetic properties of these mutants in site-specific recombination solely to the flanking sequences Δ or Δ' .

Comparison of the mutant core sequences with those of the wild-type reveals a conserved five-base sequence in the center of the core



The conservation of this region in four different mutant *att* sites is suggestive evidence for it being at least one of the critical regions necessary for some aspect of site-specific recombination. Our present data do not include a sufficient number of mutant *att* sites to define this critical region with great certainty or precision. In two *att* sites that might be regarded as analogous to those described here, the left (Δ OP') and right (POA') prophage *att* sites of a λ chromosome inserted into the *trpC* gene of *E. coli* have a six-base-pair conserved sequence located at the center of each of the core regions (C. Christie and T. Platt, personal communication). It is interesting, however, that calculations based on (i) the number of secondary *att* sites within a segment of the *E. coli* genome (33) and (ii) the number of *int*-dependent derived deletions in the nonessential region of λ (3, 31) predict a five- or six-base-pair recognition sequence to account for the apparent specificity of integrative recombination. This approximation is, of course, based upon the simplifying assumption of a random DNA sequence for the genomes of *E. coli* and λ .

It is clear that in a functional sense both the P and P' arms cannot be replaced by the corresponding Δ and Δ' sequences. One may now ask whether it is possible to replace the B and B'

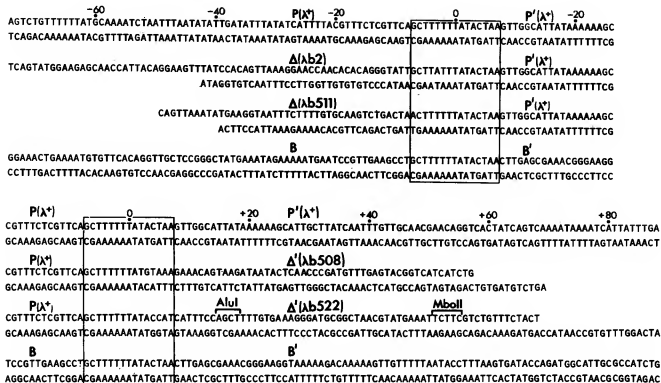


FIG. 4. The nucleotide sequence surrounding the common core of the *att* sites in λ b2, λ b511, λ b508, and λ b522. For comparison, the wild-type sequence of the P and B arms (18) have been placed above and below, respectively, the A arms of λ b2 and λ b511. Likewise, the P' and B' arms have been aligned with the two A' arms of λ b508 and λ b522. The core structures are denoted by boxed regions. Where only the sequence of one strand is shown, we were unable to obtain the complementary strand because of proximity to the 3' terminus. In these regions, sequence was confirmed by repeated sequencing of the strand shown.

arms without loss of function. In the case of the B arm, the answer is certainly yes. Although λ b2 and λ b511 do not recombine efficiently with a BOB' *att* site, they do recombine efficiently with a POB' *att* site (3, 31, 32). In fact, their genetic behavior in site-specific recombination is completely analogous to that of the BOP' *att* site; BOP' does not recombine efficiently with BOB' but does recombine efficiently with POB' (3, 12, 31). From this genetic data it appears that the A arms of λ b2 and λ b511 are equivalent to the B arm in terms of their contribution to *int*-mediated site-specific recombination—that is, without any apparent homology between the A arms and the B arm.

For the replacement of B' with Δ ' sequences, the genetic data are more variable. In this instance, *int*-dependent recombination between the POA' and BOP' *att* sites proceeds at 1/10th to 1/2 the efficiency of the analogous recombination between POB' and BOP' (3, 31). Because the range of measurable re-

combination efficiencies extends over several orders of magnitude, we do not view this as a very large reduction in efficiency for the POA' sites. Considering the absence of any apparent homology between the Δ ' sequences and B' sequences, it seems remarkable in fact that the Δ ' arms would function at all if the B' arm sequence were essential for site-specific recombination. (The relatively small reduction in recombination efficiency might well be explained by the changes found in the core structure.) Thus, our tentative interpretation of the results reported here in conjunction with prior genetic data on the behavior of these mutants in site-specific recombination lead us to favor a model similar to one suggested by Shulman *et al.* (5): Site-specific *int*-dependent recombination requires only three DNA elements—the common core region, O, the P arm, and the P' arm. It should be emphasized, however, that this model cannot be satisfactorily tested without finding, or constructing, *att* sites that have altered arms flanking a completely wild-type core region.

The interaction of these mutant *att* sites with purified Int protein presents some surprising results. Kikuchi and Nash (34, 35) have shown that, in a filter-binding assay, purified Int protein (in the absence of any other factors) binds to POP', BOP', and Δ OP' (λ b2 and λ b511) with equal efficiency and 20–30 times more efficiently than to BOB', POB', and POA' (λ b517). [According to Davis and Parkinson (17), λ b517 is probably the same as λ b508 sequenced in this paper.] These data argue strongly in favor of an important role for the P' arm in terms of Int-binding. An important role for P' was also inferred from the structure of the wild-type phage *att* site; an 11-base-pair sequence of the common core region is present in the P' arm as a perfect inverted repeat (18). It is clear from the sequences reported here that phases that have maintained the inverted repeat in P' (λ b2 and λ b511) bind Int more efficiently

	-7	0	+7
λ^*	G	C T T T T T	A T A C T A A
λ b2 (Δ P')	G	C T T A T T T	A T A C T A A
λ b511 (Δ P')	A	C T T T T T T	A T A C T A A
λ b508 (Δ P')	G	C T T T T T T	A T G T A A A
λ b522 (Δ P')	G	C T T T T T T	A T A C C A T
CONSERVED	C T T	T T T A T	A

FIG. 5. Comparison of the nucleotide sequences of the cores from λ^* , λ b2, λ b511, λ b508, and λ b522. Bases that are common to all five cores are designated "conserved."

than does λ 508 which no longer has the P' inverted repeat. Furthermore, the single base changes in the core regions of λ b2 and λ b511 do not seem to reduce the efficiency with which either of these bind Int protein in the filter assay. It should be pointed out that integrative recombination is a complex reaction and, therefore, filter binding assays of purified Int may reflect only one aspect of the role of this protein in the overall process.

We thank C. Foeller for technical assistance, S. Parkinson and F. Blattner for phage strains, M. McNamara for preparation of the manuscript, and L. Enquist, H. Nash, and R. Weisberg for critical reading of the manuscript. This work was supported by grants from the National Institutes of Health (CA11208) and The National Foundation-March of Dimes (1-543). R.H. is the recipient of a National Institutes of Health Postdoctoral Research Fellowship (GM 05269). A.L. is a Faculty Research Associate of the American Cancer Society, Inc.

1. Nash, H. (1977) *Curr. Top. Microbiol. Immunol.* 78, 171-199.
2. Kellenberger, G., Zichichi, M. L. & Weigle, J. (1961) *J. Mol. Biol.* 3, 399-408.
3. Parkinson, J. S. (1971) *J. Mol. Biol.* 56, 385-401.
4. Shulman, M. & Gottesman, M. (1973) *J. Mol. Biol.* 81, 461-482.
5. Shulman, M. J., Mizuuchi, K. & Gottesman, M. M. (1976) *Virology* 72, 13-22.
6. Campbell, A. (1962) *Adv. Genet.* 11, 101-145.
7. Zissler, J. (1967) *Virology* 31, 189.
8. Gingery, R. & Echols, H. (1967) *Proc. Natl. Acad. Sci. USA* 58, 1507-1514.
9. Gottesman, M. E. & Yarmolinsky, M. B. (1968) *J. Mol. Biol.* 31, 487-505.
10. Miller, H. I., Kikuchi, A., Nash, H., Weisberg, R. A. & Friedman, D. I. (1978) *Cold Spring Harbor Symp. Quant. Biol.* 43, in press.
11. Williams, J. K. G., Wulff, D. L. & Nash, H. A. (1977) in *DNA Insertion Elements, Plasmids and Episomes*, eds. Bukhari, A. I., Shapiro, J. & Adhya, S. (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY), pp. 357-361.
12. Gottesman, M. E. & Weisberg, R. A. (1971) in *The Bacteriophage Lambda*, ed. Hershey, A. D. (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY), pp. 113-138.
13. Kaiser, A. & Masuda, T. (1970) *J. Mol. Biol.* 47, 557-564.
14. Guarnieros, G. & Echols, H. (1970) *J. Mol. Biol.* 47, 565-574.
15. Signer, E. R., Weil, J. & Kimball, P. C. (1969) *J. Mol. Biol.* 46, 543-563.
16. Guerrini, F. (1969) *J. Mol. Biol.* 46, 523-542.
17. Davis, R. W. & Parkinson, J. S. (1971) *J. Mol. Biol.* 56, 403-423.
18. Landy, A. & Ross, W. (1977) *Science* 197, 1147-1160.
19. Robinson, L. H. & Landy, A. (1977) *Gene* 2, 1-31.
20. Marini, J. C., Weisberg, R. A. & Landy, A. (1977) *Virology* 83, 254-270.
21. Egan, J. & Landy, A. (1978) *J. Biol. Chem.* 253, 3607-3622.
22. Rigby, P. W. J., Dieckmann, M., Rhodes, C. & Berg, P. (1977) *J. Mol. Biol.* 113, 237-251.
23. Maxam, A. & Gilbert, W. (1977) *Proc. Natl. Acad. Sci. USA* 74, 560-564.
24. Sanger, F. & Coulson, A. R. (1978) *FEBS Lett.* 87, 107-110.
25. Davidson, N. & Szysbalski, W. (1971) in *The Bacteriophage Lambda*, ed. Hershey, A. D. (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY), pp. 45-82.
26. Thomas, M. & Davis, R. W. (1975) *J. Mol. Biol.* 91, 315-328.
27. Haggerty, D. M. & Schleif, R. F. (1976) *J. Virol.* 18, 659-663.
28. Wellauer, P. K., Roeder, R. H., Carroll, D., Brown, D. D., Deutch, T., Higashinakagawa, T. & David, I. (1974) *Proc. Natl. Acad. Sci. USA* 71, 2823-2827.
29. Smith, H. O. & Birnstiel, M. L. (1976) *Nucleic Acids Res.* 3, 2387-2398.
30. Davies, R. W., Schreier, P. H. & Büchel, D. E. (1977) *Nature (London)* 270, 757-760.
31. Shimada, K., Weisberg, R. A. & Gottesman, M. E. (1975) *J. Mol. Biol.* 93, 415-429.
32. Fischer-Fantuzzi, L. (1967) *Virology* 32, 18-32.
33. Boulter, J. & Lee, N. (1975) *J. Bacteriol.* 123, 1043-1054.
34. Kikuchi, V. & Nash, H. (1978) *J. Biol. Chem.*, in press.
35. Kikuchi, V. & Nash, H. (1978) *Cold Spring Harbor Symp. Quant. Biol.* 43, in press.

EXHIBIT B

NOTES

Spontaneous Deletion Mutants of the *Lactococcus lactis* Temperate Bacteriophage BK5-T and Localization of the BK5-T *attP* Site

JOHN D. BOYCE,¹ BARRIE E. DAVIDSON,¹ AND ALAN J. HILLIER^{2*}

¹Russell Grimwade School of Biochemistry and Molecular Biology, University of Melbourne, Parkville, Victoria 3052,
and Division of Food Science and Technology, Melbourne Laboratory, Commonwealth Scientific and
Industrial Research Organization, Highett, Victoria 3190,² Australia

Received 6 April 1995/Accepted 16 August 1995

Spontaneous deletion mutants of the temperate lactococcal bacteriophage BK5-T were obtained when the phage was grown vegetatively on the indicator strain *Lactococcus lactis* subsp. *cremoris* H2. One deletion mutant was unable to form stable lysogens, and analysis of this mutant led to the identification of the BK5-T *attP* site and the integrase gene (*int*). The core sequences of the BK5-T *attP* and host *attB* regions are conserved in a number of lactococcal phages and *L. lactis* strains.

BK5-T is a temperate lactococcal bacteriophage that can be induced from the lysogen *Lactococcus lactis* subsp. *cremoris* BK5 by mitomycin treatment (9). Under certain propagation conditions, BK5-T spontaneously loses the ability to lysogenize, possibly as a result of one or more deletion events (7). The availability of detailed information concerning the BK5-T genome (3, 11) prompted us to seek deletion mutants of BK5-T that had lost the ability to form stable lysogens as a strategy for identifying genes essential for the establishment and/or maintenance of lysogeny. Characterization of one of these mutants enabled us to identify and sequence the phage and host attachment sites, *attP* and *attB*, respectively. Deletion of a 536-codon open reading frame (ORF) and tandemly repeated segments within a 1,904-codon ORF in BK5-T did not affect the frequency of lysogeny, thereby eliminating the possibility that these gene products are required for lysogeny.

Deletions within the BK5-T genome during lytic propagation. The indicator strain *L. lactis* H2 (9) was infected with BK5-T.H2L (BK5-T isolated after induction of the lysogen *L. lactis* H2L [3]) and incubated until the culture lysed. The culture was centrifuged (5,000 × g) for 10 min, and the cell-free lysate was added to an appropriate volume of uninfected cells. This lysis/infection cycle was repeated 20 times, and the resulting phage was designated BK5-T.H2cyc20. Analysis of *EcoRI*-*PstI* digests of BK5-T.H2L DNA and BK5-T.H2cyc20 DNA, together with hybridization studies with *EcoRI*-a and *EcoRI*-b (11), showed that at least nine fragments (in addition to the fragments containing *cos*) in the BK5-T.H2cyc20 DNA digest were present in submolar amounts (Fig. 1). Five of these fragments were produced by specific deletions within *EcoRI*-b (11) (Fig. 1B, lane 2). *EcoRI*-b contains four perfect tandem repeats of 468 bp and a fifth incomplete tandem repeat within the large ORF1904 (3). The 7.4-kbp fragment that hybridized to *EcoRI*-b was the full-length *EcoRI*-b(P1) fragment (11), while the sizes of four other submolar fragments (6.9, 6.4, 6.0,

and 5.5 kbp) were consistent with their being produced by the loss of one, two, three, or four of the 468-bp tandem repeats, respectively. The other four submolar fragments (5.2, 4.3, 3.2, and 2.1 kbp in size) resulted from deletions within *EcoRI*-a (11) (Fig. 1C, lane 2).

Characterization of BK5-T deletion mutants. The presence of many submolar fragments in restriction digests of BK5-T.H2cyc20 DNA suggested that the BK5-T.H2cyc20 preparation comprised a mixed population. To characterize the deletions more precisely, three deletion mutants of BK5-T (BK5-T.H2Δ8, BK5-T.H2Δ10, and BK5-T.H2Δ11) were randomly selected from single plaques of BK5-T.H2cyc20. Deletions within BK5-T.H2Δ8, BK5-T.H2Δ10, and BK5-T.H2Δ11 DNA were located by restriction mapping, Southern hybridization (Fig. 1), and DNA sequencing. Each mutant contained a deletion within the tandem repeat region of ORF1904 (Fig. 2) (3), while BK5-T.H2Δ8 and BK5-T.H2Δ10 also contained deletions within *EcoRI*-a. The deletion within ORF1904 could be explained by a single crossover recombination event between two 5-bp "core sequences" of ACGGA situated at 4198 and 6070 bp (3) in BK5-T.H2Δ11 DNA or between two 8-bp homologous regions situated at 4222 and 6094 bp (3) in BK5-T.H2Δ8 DNA (Fig. 2). Each of these deletions removed 1,872 bp of DNA, equivalent to four of the 468-bp tandem repeats identified previously (3). These deletions shortened ORF1904 by 624 codons without changing the reading phase. BK5-T.H2Δ10 phage comprised a mixed population of phages carrying deletions within ORF1904 corresponding to the loss of two, three, or four 468-bp repeats. The precise endpoints of these deletions could not be determined. Each phage deletion mutant was able to propagate vegetatively on *L. lactis* H2, yielding slightly larger plaques (1 to 1.5 mm) than did BK5-T.H2L (0.5 to 1 mm).

FASTA (15) comparison of the ORF1904 amino acid sequence with all protein sequences in GenBank showed significant homology with a number of proteins from the collagen family. This homology was centered around the repeated Gly-X-Y motif, found 64 times in ORF1904. FASTA analysis also indicated similarity between ORF1904 and an unidentified ORF (ORF35) from the lactococcal bacteriophage bIL67 (19) and less similarity to the lactococcal lytic bacteriophage US3

* Corresponding author. Mailing address: Division of Food Science and Technology, Melbourne Laboratory, Commonwealth Scientific and Industrial Research Organization, P.O. Box 20, Highett, Victoria 3190, Australia. Phone: 61 39344 5915. Fax: 61 39347 7730. Electronic mail address: alanh@mel.dlbc.esiro.au.

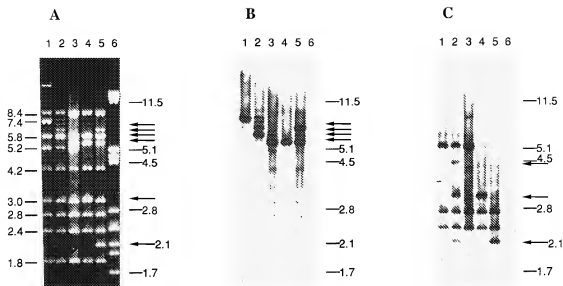


FIG. 1. Electrophoretic separation of *EcoRI-PstI* restriction digests of DNA from BK5-T.H2L and deletion mutants. (A) BK5-T DNA was digested with both *EcoRI* and *PstI*, and the resulting fragments were electrophoresed through a 0.8% (wt/vol) agarose gel and stained with ethidium bromide. The DNA was from BK5-T.H2L (lane 1), BK5-T.H2cyc20 (lane 2), BK5-T.H2Δ11 (lane 3), BK5-T.H2Δ8 (lane 4), BK5-T.H2Δ10 (lane 5), and λ DNA digested with *PstI* (lane 6). (B) Southern blot of the gel shown in panel A after hybridization with 32 P-labeled *EcoRI*-b (11). (C) Southern blot of the gel shown in panel A after hybridization with 32 P-labeled *EcoRI*-a (11). Numbers on the right of each panel indicate the sizes (in kilobase pairs) of fragments from a *PstI* digest of a DNA, and arrows show positions of fragments present in the digests of mutant phages which were not present in the digest of BK5-T.H2L. Numbers on the left of panel A indicate the sizes (in kilobase pairs) of BK5-T.H2L fragments.

lytic enzyme (16) and the *Bacillus subtilis* xylose isomerase (22). Analysis by COMPARE (18) indicated similarity between ORF1904 and numerous proteins involved in binding and/or degradation of cell wall glycoproteins. These sequence similarities and the observation that a number of cell wall-lytic en-

zymes contain repeated sequence motifs (5, 8, 10, 12) suggest that ORF1904 may be involved in cell lysis during the lytic cycle of BK5-T or in cell wall hydrolysis to enable phage DNA injection.

The genomes of BK5-T.H2Δ8 and BK5-T.H2Δ10 also con-

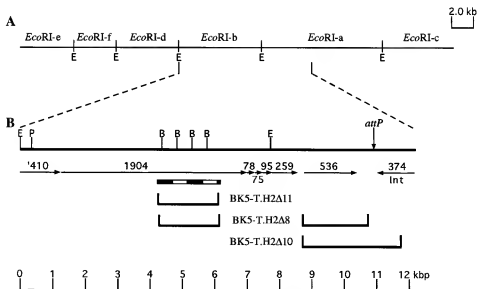


FIG. 2. Physical map of the BK5-T genome. (A) *EcoRI* (E) restriction map of the BK5-T genome. Restriction fragments are named as described previously (11). The ends of the map correspond to the *cos* ends of the phage DNA (3). (B) Physical map of *EcoRI*-b and part of *EcoRI*-a showing the locations of all *EcoRI* (E), *PstI* (P), and *BglI* (B) sites and the positions of relevant ORFs as determined by nucleotide sequence analysis (3). Horizontal arrows indicate the orientation of each ORF, and adjacent numbers show the number of codons in each ORF from the first ATG codon to the first stop codon. ORF410 is open at the 5' end. The positions of the tandem duplications are indicated by solid and open rectangles. *attP* is the phage attachment site. The positions of the deletions within the genomes of BK5-T.H2Δ11, BK5-T.H2Δ8 (two separate deletions), and BK5-T.H2Δ10 phage are bracketed. Since BK5-T.H2Δ10 was a mixed population of phages carrying various deletions within *EcoRI*-b, only the deletion within *EcoRI*-a has been shown.

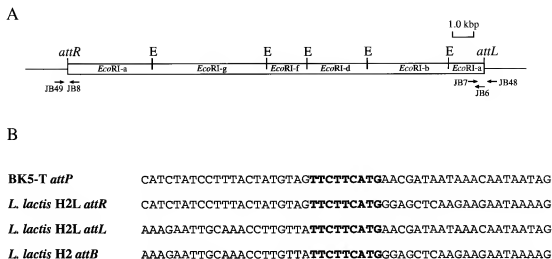


FIG. 3. Nucleotide sequence of the att regions of BK5-T, *L. lactis* H2, and *L. lactis* H2L. (A) Physical map of the BK5-T prophage. BK5-T is designated by the open rectangle, while the *L. lactis* chromosomal DNA is indicated by the horizontal line. The positions of oligonucleotides used to amplify the *attL* and *attR* regions are shown by arrows, with the arrowheads designating the 3' end of each oligonucleotide. (B) Nucleotide sequences of *attP*, *attR*, *attL*, and *attB*. The designation of *attL* and *attR* was in accordance with the system of Lillehaug and Birkehead (13). The nucleotide sequence of each region was determined as described in the text. The 9-nucleotide core sequence 5'-TTCTTCATG-3' present in all att regions is shown in boldface type. The sequence of the *attL*, *attR*, and *attB* regions was determined in only one strand.

tained deletions within *EcoRI*-a (Fig. 2). These deletions resulted from a single-crossover recombination event between two 11-bp core sequences of TTTTITTTGTTT situated at 8677 and 10701 bp in BK5-T.H2Δ8 DNA (3) (Fig. 2) or between two 6-bp core sequences of GTGTGTT situated at 8674 and 11737 bp in BK5-T.H2Δ10 DNA (3) (Fig. 2). Both of these deletions removed ORF536, while the BK5-T.H2Δ10 deletion also removed the ORF374-to-ORF536 intergenic region and 256 codons from the C-terminal end of ORF374 (Fig. 2).

The phage deletion mutants were tested for their ability to form lysogens on *L. lactis* H2, BK5-T.H2Δ11 and BK5-T.H2Δ8 remained lysogenic, but BK5-T.H2Δ10 was not. This indicated that neither the full-length ORF1904 nor ORF536 (Fig. 2) is essential for vegetative growth or the formation and maintenance of stable lysogens. Coliphage λ also contains genes that encode products that are not essential for normal propagation (6). The function of the ORF536 protein is unknown, and its amino acid sequence exhibits no significant homology with any prokaryotic proteins in the GenBank database. The inability of BK5-T.H2Δ10 to form lysogens indicates that ORF374 and/or the ORF536-to-ORF374 intergenic DNA is essential for the establishment and/or maintenance of lysogeny.

Localization of the BK5-T attachment site. Since BK5-T.H2Δ10 contained deletions within the *EcoRI* a fragment, previously shown to contain *attP* (11), it was decided to locate *attP* more precisely to determine whether loss of this feature was related to the nonlysogenic nature of BK5-T.H2Δ10. Subfragments of BK5-T *EcoRI*-a were used to probe Southern blots of *EcoRI* digests of chromosomal DNA from the BK5-T lysogens *L. lactis* BK5 and H2L (data not shown). BK5-T subfragments spanning *attP* would hybridize to two chromosomal fragments, whereas subfragments which did not contain *attP* would hybridize to only one fragment. By using this approach, *attP* was localized between 10769 and 10999 bp (3) and the phage/host junctions were shown to be in chromosomal *XbaI* fragments of 90 and 18 kbp in *L. lactis* BK5 and H2L (data not shown).

The sequences of these phage/host junctions (*attL* and *attR*) in *L. lactis* H2L were then determined (Fig. 3B). For *attL*, this

was done by sequencing a 900-bp PCR fragment, obtained by inverse PCR (14) with JB6 and JB7 as primers and a ligated *SspI* digest of the gel-purified 90-kbp *XbaI* chromosomal fragment as the template. The obtained sequence revealed 100% sequence identity between the 21 bp of host DNA adjacent to *attL* in BK5-T and in φLC3, another *L. lactis* temperate phage (13). The existence of this identity and of the 97% identity between the 1,627 bp of DNA surrounding *attP* of BK5-T and φLC3 (13) enabled us to use the nucleotide sequence surrounding *attB* in the φLC3 lysogen *L. lactis* IMN-C18 (13) to design PCR primers that amplified *L. lactis* H2 *attB* (JB48 and JB49 [Table 1]) and *L. lactis* H2L *attR* (JB8 and JB49 [Table 1]) and provided DNA for sequencing.

Comparison of the DNA sequences of the *attP*, *attL*, *attR*, and *attB* regions (Fig. 3B) identified a common 9-bp core sequence, 5'-TTCTTCATG-3' (bases 10882 to 10874) (3), within which recombination between BK5-T and the host genome is likely to occur. Thus, BK5-T *attP* is located in a region of the BK5-T genome that was deleted in BK5-T.H2Δ10 but not in BK5-T.H2Δ11 or BK5-T.H2Δ8 (Fig. 2). The deletion of *attP* in BK5-T.H2Δ10 provides an explanation for the inability of this phage to form stable lysogens. This observation is of particular interest in an industrial context, since to our knowledge it is the first demonstration of spontaneous mutations

TABLE 1. Oligonucleotides used in this investigation^a

Number	Oligonucleotide sequence (5'-3')	Region of BK5-T ^b
JB6	GATCAITAGGAATACTCCCC	10012-9994
JB7	GATCGACATGGGAGAAGGTAAAGG	10259-10281
JB8	CACACAGCAAACTATATCC	11051-11032
JB48	TGTTAAAGCAGGAATCAAAGG	Comp ^c
JB49	AATACCTAAGCACACGAAGGT	Comp

^a Oligonucleotides were synthesized in an Applied Biosystems model 381 DNA synthesizer.

^b Sequence numbers refer to the nucleotide sequence of BK5-T as determined previously (3).

^c Comp, complementary to *L. lactis* IMN-C18 genomic DNA (13).

9. Huggins, A. R., and W. E. Sandine. 1977. Incidence and properties of temperate bacteriophages induced from lactic streptococci. *Appl. Environ. Microbiol.* 33:184-191.
10. Kawazu, T., Y. Nakanishi, N. Uozumi, T. Sasaki, H. Yamagata, N. Tsukagoshi, and S. Ueda. 1987. Cloning and nucleotide sequence of the gene coding for enzymatically active fragments of the *Bacillus polymyxa* β -amylase. *J. Bacteriol.* 169:1564-1570.
11. Lakshmidivi, G., B. E. Davidson, and A. J. Hillier. 1988. Circular permutation of the genome of a temperate bacteriophage from *Streptococcus cremoris* BK5. *Appl. Environ. Microbiol.* 54:1039-1045.
12. Lemaire, M., and P. Beguin. 1993. Nucleotide sequence of the *celG* gene of *Clostridium thermocellum* and characterization of its product, endoglucanase CelG. *J. Bacteriol.* 175:3353-3360.
13. Lillehaug, D., and N. Birkeland. 1993. Characterization of genetic elements required for site-specific integration of the temperate lactococcal bacteriophage ϕ LC3 and construction of integration-negative ϕ LC3 mutants. *J. Bacteriol.* 175:1745-1755.
14. Ochman, H., A. S. Gerber, and D. L. Hartl. 1988. Genetic applications of an inverse polymerase chain reaction. *Genetics* 120:621-623.
15. Pearson, W. R., and D. J. Lipman. 1988. Improved tools for biological sequence comparison. *Proc. Natl. Acad. Sci. USA* 85:2444-2448.
16. Platteaux, C., and W. M. de Vos. 1992. Location, characterization and expression of lytic enzyme-encoding gene, *htrA*, of *Lactococcus lactis* bacteriophage ϕ US3. *Gene* 118:115-120.
17. Raya, R. R., C. Frenaux, G. L. De Antoni, and T. R. Klaenhammer. 1992. Site-specific integration of the temperate bacteriophage ϕ adh into the *Lactobacillus gasseri* chromosome and molecular characterization of the phage (*attP*) and bacterial (*attB*) attachment sites. *J. Bacteriol.* 174:5584-5592.
18. Reiser, A. H., and N. H. Westwood. 1982. Assessment of similarities of pairs and groups of proteins using transformed amino-acid-residue data. *J. Mol. Evol.* 18:240-250.
19. Schouler, C., S. D. Ehrlich, and M. Chopin. 1994. Sequence and organization of the lactococcal prolate-headed ϕ IL67 phage genome. *Microbiology* 140: 3061-3069.
20. van de Guchte, M., C. Daly, G. F. Fitzgerald, and E. K. Arendt. 1994. Identification of *int* and *attP* on the genome of lactococcal bacteriophage Tuc2009 and their use for site-specific plasmid integration in the chromosome of Tuc2009-resistant *Lactococcus lactis* MG1363. *Appl. Environ. Microbiol.* 60:2324-2329.
21. van de Guchte, M., C. Daly, G. F. Fitzgerald, and E. K. Arendt. 1994. Identification of the putative repressor-encoding gene *cI* of the temperate lactococcal bacteriophage Tuc2009. *Gene* 144:93-95.
22. Wilhelm, M., and C. P. Hollenberg. 1985. Nucleotide sequence of the *Bacillus subtilis* xylose isomerase gene: extensive homology between the *Bacillus* and *Escherichia coli* enzyme. *Nucleic Acids Res.* 13:5717-5722.

EXHIBIT C

Detection of Homology to the Beta Bacteriophage Integration Site in a Wide Variety of *Corynebacterium* spp.

NICHOLAS CIANCIOOTTO,¹ RINO RAPPUOLI,² AND NEAL GROMAN^{1*}

Department of Microbiology and Immunology, School of Medicine, University of Washington, Seattle, Washington 98195,¹ and Scavo Research Center, Siena, Italy²

Received 10 April 1986/Accepted 1 July 1986

In toxigenic conversion of *Corynebacterium diphtheriae* C7, β bacteriophage DNA integrates into either of two chromosomal attachment sites, *attB1* or *attB2*. These *attB* sites share a 96-base-pair sequence with the *attP* sites of β -related phages. The distribution of *attB*-related sites in other species of *Corynebacterium* was assessed by hybridization with a DNA probe containing both *attB* sites of the C7 strain and a second DNA probe containing the *attP* site of a β -related phage. All but one of the 15 *C. diphtheriae* strains tested, regardless of origin or colonial type, contained at least two *Bam*HI fragments that hybridized strongly to both of these probes under conditions of high stringency. Strains of *C. ulcerans* and *C. pseudotuberculosis*, species in which conversion to toxigeny has also been demonstrated, also had one or two hybridizing *Bam*HI fragments. The functionality of these sites as integration sites was demonstrated by isolating lysogens of all three species following single infection with one or more β -related phages. As predicted, following lysogenization one of the DNA fragments that had exhibited homology with the *attB1-attB2* probe was replaced by two hybridizing fragments. Other species of *Corynebacterium*, including pathogens and nonpathogens from animals, plant pathogens, and soil isolates also carried at least one *Bam*HI fragment that hybridized with the *attB1-attB2* and *attP* probes. The data indicate that sequences homologous to the β phage integration sites in *C. diphtheriae* have been conserved in members of the genus *Corynebacterium*.

Nontoxigenic strains of *Corynebacterium diphtheriae* are converted to toxigeny by members of the β family of corynebacteriophage (6). In addition diphtheria toxin-positive strains of *C. ulcerans* and *C. pseudotuberculosis* may be produced by phage conversion or may already carry *tox*⁺ β -related phages (1, 5, 6, 9). Genetic evidence and Southern hybridization experiments suggest that converting phages integrate into the bacterial chromosome by a process analogous to lambda lysogenization of *Escherichia coli* (2, 8, 10). Presumably, site-specific recombination occurs between a phage attachment site (*attP*) and a bacterial attachment site (*attB*). It has been shown that the *C. diphtheriae* genome contains two *attB* sites, *attB1* and *attB2* (13), and recently Rappuoli and Ratti (14) cloned and characterized the region of the *C. diphtheriae* C7 chromosome containing these sites. The *attB* sites are 2.25 kilobases apart but share a 96-base pair (*bp*) sequence which is also found in the *attP* sites of the closely related β , γ , and ω phages [1, 2, 12; G. Ratti and R. Rappuoli, J. Cell. Biochem., 7B(Suppl.):155, 1983; personal communication]. Since lysogenic conversion of various strains of *C. diphtheriae*, *C. ulcerans*, and *C. pseudotuberculosis* requires integration of a β -related phage, it was of interest to determine whether members of all three species contained sites similar to the *attB* sites of C7. In the present study, using *attB* and *attP* probes, sites homologous to those in C7 were detected in all strains of the three species examined. They were also detected in other pathogenic and nonpathogenic *Corynebacterium* spp. not presently known to be susceptible to or carriers of β -related phages.

MATERIALS AND METHODS

Bacteria, phages, and media. All of the bacterial strains used were from our culture collection. The strains are

mentioned specifically in the text or in figures, and their origins are listed in Table 1. In addition, phage 782-carrying lysogens of *C. ulcerans* 690, 712, and 872 and *C. pseudotuberculosis* 766 were isolated, as well as a C7 lysogen of phage 76 (9) and a *C. renale* 412 lysogen of phage RP28 (7). All of the other lysogens used and their sources and the sources of *C. diphtheriae* phages β , γ , π , ω , δ , κ , ρ , and 782 and *C. ulcerans* phages 731, 761, 876, a, and h were previously described (1, 3). Phage 782 is a *tox*⁰ β -related phage, i.e., one that fails to hybridize with probes specific for the diphtheria toxin gene *tox*.

Bacteria were grown in tryptose-yeast extract broth containing tryptose (10 g), yeast extract (5 g), and NaCl (5 g) per liter or in heart infusion broth (Difco Laboratories, Detroit, Mich.). The agar media contained 1.5% agar per liter.

DNA probes. The plasmid A634 contains the *attB1* and *attB2* region of the *C. diphtheriae* C7 chromosome in a pUC8 vector (14). This cloned insert, which will be referred to as the *attB1-attB2* region and probe is diagrammed in Fig. 1. As will be seen, this probe is probably not exclusively specific for *attB*-related sites; nevertheless, it proved useful in screening for such sites. An *Hinc*II fragment of A634 that served as a probe to test for DNA flanking the *attB* sites and an *attP* probe obtained by digesting plasmid T30 DNA with *Acc*I and *Eco*RI are also diagrammed. Plasmid T30 was isolated at the Scavo Research Center, Siena, Italy, which supplied the DNA for our probe. It contains the *attP-tox* region of ω phage. It has been shown that ω phage is closely related to converting phage β (12). They have homologous *attP* sites and integrate into the same *attB* sites (13). The diphtheria toxin probe specific for the B fragment coding region was previously described (4). Bacterial and plasmid DNAs were extracted as previously described (4).

Analytical procedures. The methods used for restriction enzyme digests, agarose gel electrophoresis, nick translation, dot blot and Southern hybridizations, autoradiography,

* Corresponding author.

TABLE 1. Bacterial strains

Species	Strain ^{a,b}	Source or reference ^c
<i>C. diphtheriae</i> (8) ^d	C7 (770) S1013 (8026) S1014 (8028) S1015 (8032) S1016 (19409/3984) S601	V. J. Freeman: 2 ATCC ATCC ATCC ATCC; NCTC Seattle, Wash.; 16
" <i>C. diphtheriae</i> (belfanti)"	820 (1030)	Romania; 9
<i>C. ulcerans</i> (12)	690 (9015) 712 (51169) 740 (1613/50) 751 (378) 755 (9304) 872 (A238) S1019 (19410)	ATCC; Albany, N.Y.; 5 Albany, N.Y.; 5 Norway; 5 Wales; 16 Romania; 9 Romania; 9 ATCC
<i>C. pseudotuberculosis</i> (2)	766 (992) 769 (211) 771 (1113) 411 (28) 412 (8)	Romania; 9 Romania; 9 Romania; 9 Japan; 7 Japan; 7
<i>C. renale</i>	S1001 (9775) S1002 (231)	NCTC NCTC
<i>C. xerosis</i> (4)	S1041 (10700)	ATCC
<i>C. pseudodiphtheriticum</i> (2)	S1003 (10288)	NCTC
<i>C. minutissimum</i> (1)	S1017 (764)	NCTC
" <i>C. flavidum</i> (strictum)"	S1018 (10673) S1020 (7715) S1023 (363) S1046 (6939) S1080 (13032)	NCTC ATCC NCPBP ATCC ATCC

^a The strain designations include our stock number followed in parentheses by the stock number of the source laboratory.

^b All strains listed were positive in hybridization tests with the *attB1-attB2* probe. Similarly, all strains tested with the *attP* probe, which included all of the strains listed except S1013, S1014, 755, S1019, and 771, were also positive with that probe.

^c ATCC, American Type Culture Collection, Rockville, Md.; NCTC, National Culture Type Collection, London, England; NCPBP, National Collection of Plant Pathogenic Bacteria, Harpenden, England.

^d The numbers in parentheses immediately following the species name indicate the number of strains tested by hybridization with the *attB1-attB2* probe in addition to those listed. All of the strains tested were positive with this probe.

and isolation of DNA fragments from agarose gels were previously described (2, 4). Stringency conditions for all hybridizations permitted a 10 to 12% bp mismatch.

All bacterial strains were initially screened for the presence of *attB* sites in dot blot hybridizations with the *attB1-attB2* probe. Southern blots were then performed on restriction enzyme digests of the DNAs of dot blot-positive strains.

RESULTS

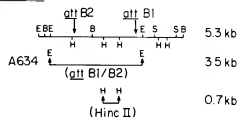
Presence of *attB1-attB2* in strains of *C. diphtheriae*. It was previously shown that *tox*⁺ β and related α phage and *tox*⁺ γ phage integrated into the *attB1-attB2*-containing region of *C. diphtheriae* C7 (13, 14). To determine whether this C7

fragment was consistently used as a site of integration, we examined C7 lysogens carrying other phages. Integration was detected by the disappearance of one and the appearance of two new fragments hybridizing with the *attB1-attB2* probe after lysogenization of C7. Using this criterion, we found that the β -related *tox*⁺ phages π , δ , and 76 (data not shown) and the β -related *tox*⁺ phage 782 (see Fig. 3) all integrated into a *Bam*HI fragment that hybridized with the probe, whereas phages κ and ρ , which are unrelated to β or to each other, did not (1). Thus, all β -related phages tested integrated preferentially into the region defined by the *attB1-attB2* probe. Though these phages displayed a preference for *attB2* in our study, this was not the case in a previous study in which the same strain of *C. diphtheriae* was used (13).

Fifteen strains of *C. diphtheriae*, including representatives of the three colonial types and the subspecies of "*C. diphtheriae* (belfanti)," were next tested for the presence of DNA homologous to the *attB1-attB2* probe. The strains chosen varied in the time and place of their isolation, DNA restriction pattern, and presence of the *tox* gene or β -related DNA. Of the 15 tested, 6 of which are listed in Table 1 and shown in Fig. 2, all but one (lane B) contained at least two *Bam*HI fragments which hybridized strongly with the *attB1-attB2* probe. One of these fragments comigrated with the C7 *attB2*-containing fragment, whereas the size of the second varied (Fig. 2). In strains known to carry β phage-related DNA, a third hybridizing band was observed (Fig. 2, lanes C and D). This would be expected if the phage integrated at one of the *attB* sites. As expected, the third band was weak when the phage inserted into the *attB1* site since the probe extends only a short distance to one side of that site (Fig. 1).

Presence of *attB1-attB2* in strains of *C. ulcerans* and *C. pseudotuberculosis*. In addition to *C. diphtheriae*, strains of

(A) *C. diphtheriae* C7



(B) Phage α

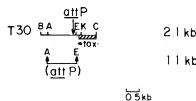


FIG. 1. DNA probes. (A) Restriction map of the region of the *C. diphtheriae* C7 chromosome containing *attB1* and *attB2* (14). Plasmid A634 contains both *attB* sites on an *Eco*RI fragment cloned into pUC8, which is referred to as the *attB1-attB2* probe. A 0.7-kilobase *Hinc*II fragment called the *Hinc*II probe was used to probe for DNA homologous to sequences adjacent to the *attB* sites. (B) Restriction map of the *attP* region of phage α cloned in pUC8 to form plasmid T30. This insert also contains the carboxyl end of the *tox* gene for diphtheria toxin. An *Acc*I-*Eco*RI fragment of T30, which contains *attP* but not *tox*, served as the *attP* probe. Abbreviations: A, *Acc*I; B, *Bam*HI; C, *Clal*; E, *Eco*RI; H, *Hinc*II; K, *Kpn*I; S, *Sal*I; kb, kilobases.

C. ulcerans and *C. pseudotuberculosis* can also be converted to toxinogeny by β -related *tox*⁺ phages (6, 9). To examine these species for sites homologous to the *attB* sites in C7, Southern blots of *Bam*HI digests of 18 strains of *C. ulcerans* and 6 of *C. pseudotuberculosis* were hybridized with the *attB1-attB2* probe. Hybridizations with a representative sample of strains of each species (listed in Table 1) are given in Fig. 3. The majority of *C. ulcerans*, such as 712 (lane C), had one hybridizing fragment; however, *tox*⁺ strains 690 (lane E) and 740 and 751 (data not shown) had two hybridizing fragments. Strain 751, as well as four other strains with two hybridizing fragments, all contained β -related DNA in excess of the *attB*-related sites and *tox*, whereas strains 690 and 740 did not (6). Five of the *C. pseudotuberculosis* strains had two hybridizing *Bam*HI fragments and the sixth, *tox*⁺ strain 766 (lane G), had three. Overall, strains of both species, regardless of their origin or the presence of the *tox* gene or β -related DNA in their genomes, had one or more *Bam*HI fragments that hybridized with the probe. The size of the hybridizing fragment(s) varied, except in several cases in which strains with similar overall DNA restriction patterns had similar-sized fragments. In all cases, the degree of hybridization with DNAs from *C. ulcerans* and *C. pseudotuberculosis* was significantly weaker than with DNAs from strains of *C. diphtheriae*. Nevertheless, the results suggested that these species contained *attB* sites homologous to those in *C. diphtheriae*.

If the *attB1-attB2*-hybridizing fragments contain functional phage integration sites, then on lysogenization with β or β -related phage two new hybridizing fragments should

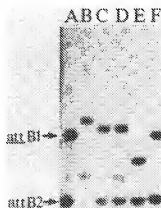


FIG. 2. Detection of *attB1-attB2* homology in *C. diphtheriae*. Purified bacterial DNA was digested with *Bam*HI and electrophoresed in 1% agarose. A Southern blot was made, hybridized with ³²P-labeled A634 DNA, and autoradiographed. Lanes: A, C7 *mitis*; B, S1013 *mitis*; C, S1014 *gravis*; D, S1016 *gravis*; E, S1015 *intermedius*; F, S601 *mitis*. All strains contain β -phase-related DNA except C7 and S601. The arrows identify the *attB1*- and *attB2*-containing fragments of C7 in lane A.

ABCDEFGH

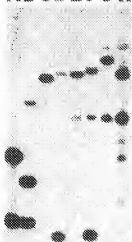


FIG. 3. Detection of *attB1-attB2* homology in *C. ulcerans* and *C. pseudotuberculosis*. Purified bacterial DNA digested with *Bam*HI was processed as described in Fig. 2. *C. diphtheriae* (lanes): A, C7; B, C7(782). *C. ulcerans* (lanes): C, 712; D, 712(782); E, 690; F, 690(782). *C. pseudotuberculosis* (lanes): G, 766; H, 766(782).

appear while one disappears. We tested this prediction in a number of instances with strains sensitive to β -related phages. When *C. ulcerans* 712 was lysogenized with *C. diphtheriae* phage 782, the result was as predicted for integration (Fig. 3, lanes C and D), as were those obtained when strain 712 was lysogenized with the β -related *tox*⁺ *C. ulcerans* phages 731, 761, or 876. Similarly, the *attB1-attB2*-hybridizing fragments provided integration sites when *C. ulcerans* 690 was lysogenized by phage 782 (Fig. 3, lanes E and F), strain 755 was lysogenized by phage ω , and strain 872 was lysogenized by *C. ulcerans* phages 731, 761, and h and *C. diphtheriae* phage 782. Surprisingly, when β -related phage 782 lysogenized *C. ulcerans* 740 and 751, integration into the *attB1-attB2* hybridizing fragment was not detected. Not surprisingly, *C. ulcerans* phage a, which is unrelated to β , did not integrate into the *attB1-attB2* region of either strain 712 or 872. The *attB1-attB2*-hybridizing fragments of *C. pseudotuberculosis* 766, 769, and 771 also served as integration sites for β -related phages 782, 76, and ω , respectively (Fig. 3, lanes G and H). Thus, in all three species, *attB1-attB2*-related regions contained functional *attB* sites for members of the β family of phage.

Hybridization of *C. diphtheriae*, *C. ulcerans*, and *C. pseudotuberculosis* with an *attP* probe. The *attB1-attB2* probe is probably not exclusively specific for *attB*, i.e., the sequences required for phage integration, since it contains about 3 kilobases of C7 DNA. Thus, strains lacking the *attB* sequence might also hybridize with that probe. To resolve this ambiguity, we used an *attP* probe derived from ω phage, and thus free of non-*attB* sequences from C7 DNA, in hybridizations. The *attP* site of ω phage shares a 96-bp sequence with the *attP* site of β and γ and the *attB1* and *attB2* sites in C7. There is no other sequence homology

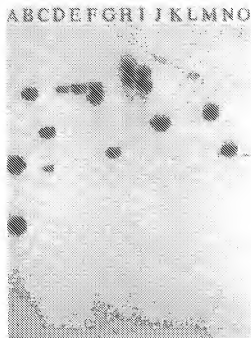


FIG. 4. Detection of *attB1-attB2* homology in *Corynebacterium* species. Purified bacterial DNA digested with *Bam*HI was processed as described in Fig. 2. Lanes: A, *C. diphtheriae* C7; B, *C. ulcerans* 712; C, *C. pseudotuberculosis* 769; D, *C. renale* 412; E, *C. renale* 412(RP28); F, *C. xerosis* S1001; G, *C. pseudodiphtheriticum* S1002; H, *C. minutissimum* S1003; I, "*C. flavidum*" S1017; J, *C. bovis* S1020; K, *C. hoagii* S1018; L, *C. betae* S1023; M, *C. glutamicum* 13032; N, *C. pseudodiphtheriticum* S1041; O, *C. equi* S1046.

between the *attP* probe and C7 *attB* region (G. Ratti, personal communication). *Bam*HI digests of DNAs from five strains of *C. diphtheriae*, five of *C. ulcerans*, and two of *C. pseudotuberculosis* were hybridized with the ω -derived *attP* probe diagramed in Fig. 1. All 12 strains hybridized with the *attP* probe and at a similar intensity with the same fragments that hybridized with the *attB1-attB2* probe (data not shown). In fact, all strains that were positive with the *attB1-attB2* probe and were tested were also positive with the *attP* probe (Table 1, footnote b).

Presence of *attB*-related sites in other *Corynebacterium* species. The presence of sites homologous to *attB1-attB2* in other *Corynebacterium* species, as well as organisms in other genera, was assessed in hybridization experiments (Fig. 4; Table 1). All *Corynebacterium* species tested hybridized with the probe. Some strains, including those of *C. xerosis*, *C. minutissimum*, "*C. flavidum*", *C. bovis*, and *C. glutamicum*, had two hybridizing *Bam*HI fragments, and those of *C. renale* and *C. pseudodiphtheriticum* had one. All of these hybridized with an intensity similar to that seen with strains of *C. ulcerans* and *C. pseudotuberculosis* but not as strong as that seen with strains of *C. diphtheriae*. Strains of *C. hoagii*, *C. betae*, and *C. equi* had one fragment, and these produced weaker hybridizations than any of the other species. When the *attP* probe was used (data not shown), it hybridized in all cases with the same *Bam*HI fragments as the *attB1-attB2* probe, but with this probe all *Corynebacterium* spp., including *C. diphtheriae*, hybridized with the same intensity, except for *C. hoagii*, *C. betae*, and *C. equi*, which produced weaker hybridizations. In other experiments (data not shown), individual strains of *Brevibacterium*

epidermidis, "*Propionibacterium aerobicum*", and five unclassified aerobic skin coryneforms exhibited one or more *Bam*HI fragments that hybridized weakly with both the *attB1-attB2* and *attP* probes. In probes with *attB1-attB2* alone, 10 strains of *C. fascians* hybridized weakly, whereas one strain of "*C. genitalium*" produced a strong reaction. However, strains of *Listeria monocytogenes*, *Staphylococcus aureus*, *Clostridium perfringens*, and *Pseudomonas aeruginosa* did not exhibit homology with either probe.

Hybridization of *Corynebacterium* spp. with DNA adjacent to the *attB* site. The observation that all *C. diphtheriae* DNAs hybridized more intensely with the *attB1-attB2* probe than did DNAs of other *Corynebacterium* species suggested that the DNA sequences flanking *attB1*, *attB2*, or both might be specific for *C. diphtheriae*. To examine this, an *Hinc*II fragment internal to the probe and located between the two *attB* sites (Fig. 1) was isolated, labeled, and hybridized with *Bam*HI digests of various *Corynebacterium* species (Fig. 5). This internal probe hybridized strongly to one of the *attB1-attB2*-hybridizing fragments of the *C. diphtheriae* strains (lanes A to E) but only weakly to those of *C. ulcerans* and *C. pseudotuberculosis* (lanes F to K), the latter becoming visible only after exposure periods days longer than those required to produce strong positive results with strains of *C. diphtheriae*. There was no detectable hybridization with the DNAs of *C. renale* or *C. glutamicum* even after prolonged exposure.

DISCUSSION

We showed that sites with homology to the *attB* phage integration sites utilized by β and β -related cory-

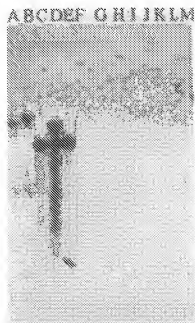


FIG. 5. Hybridizations with DNA flanking the C7 *attB* sites. Purified bacterial DNA digested with *Bam*HI was processed as described in the legend to Fig. 2. The probe was the *Hinc*II fragment of A634 as described in the legend to Fig. 1. *C. diphtheriae* (lanes): A, C7; B, S1016; C, S1015; D, S601; E, S20. *C. ulcerans* (lanes): F, 712; G, 740; H, 751; I, 872. *C. pseudotuberculosis* (lanes): J, 766; K, 769. Lane L, *C. renale* 411. Lane M, *C. glutamicum* 13032. The C7 fragment in lane A carried the *attB1* site. The DNA in lane D was partially degraded.

neobacteriophages in *C. diphtheriae* C7 were present in all of the *Corynebacterium* species and related organisms examined. The organisms carrying these sites included human and other animal pathogens, members of the normal human flora, and organisms isolated from soil and plants. Some of these organisms are known hosts for phages of the β family, but in some this relationship has not been detected and for many has yet to be examined. Where tested and testable, specifically in strains of *C. diphtheriae*, *C. ulcerans*, and *C. pseudotuberculosis*, *attB*-related sites were the site of integration for one or more members of the β family. The single exception was the failure of *tox*⁺ β -related phage 782 to integrate into *attB*-related sites in two strains of *C. ulcerans* even though it did so in others.

In contrast to the extensive information available on the requirements for the integration of lambda phage in *E. coli* (17), very little is known about the specificity requirements for the integration of members of the β phage family into *Corynebacterium* species. The data suggest that the *att* sites in both the phage and bacterium contain a common sequence of 96 bp, far larger than the 15-bp common core found in the lambda-*E. coli* system, though it is not known whether all of the 96 bp are essential to the integration process. In addition to a common core, integration of lambda depends on the presence of fairly large and specific flanking sequences outside the phage common core and, to a smaller extent, outside that of the host. The observations in this paper and those by Rappuoli et al. (13) on the integration of β -related phages and our own observations with phage 782, whose *attP*-containing *Bam*HI fragment showed no detectable homology with that of β (3), suggest that very little phage sequence specificity may be required for integration outside the 96-bp sequence shared with the bacterial genome. In addition, the weaker hybridizations we obtained with the *attB1-attB2* probe in species other than *C. diphtheriae*, in which β -related phages integrate, i.e., *C. ulcerans* and *C. pseudotuberculosis*, and their strong hybridization with the *attP* probe equivalent to that seen with *C. diphtheriae*, suggest that large flanking sequences on the bacterial chromosome are not required for integration either. Nevertheless, much remains to be learned in detail about site-specific recombination in the corynebacterial system. With respect to the anomalous behavior of phage 782 noted above, it can be argued generally that mutation either in the common core sequence of the host or an unfulfilled specificity requirement of flanking sequences prevented integration in this case. However, that alone would not explain our failure to detect the actual site of phage 782 integration with the *attP* probe, assuming that it was similar to the primary site. Additional explanations would have to be devised to account for this discrepancy. It is worth noting that, though we previously failed to detect any homology between the *attP*-containing *Bam*HI fragments of phage 782 and β phage (3), we were able to do so with the smaller, more specific *attP* probe used in the present study.

The number of *attB1-attB2*-hybridizing fragments found in various species of *Corynebacterium* differs and is a function of the intrinsic number of *attB* sites that each strain carries, variation in the *Bam*HI restriction sites which might prevent identification of two or more contiguous sites, and the presence of β -related phages. Integration of phage into an unoccupied site, in theory, yields two sites with the identical sequence. In all of the cases tested, we found that lysogenization into the *attB* site resulted in just such an increase in hybridizing sites. Strains of *C. diphtheriae*, including some not carrying β -related phages, had at least

two sites, whereas most strains of *C. ulcerans* had only one site. It has been suggested (11, 14) that the presence of two sites in *C. diphtheriae* may have been due to duplication of the *attB* site and presumably enhanced the possibility of increased diphtheria toxin production by permitting integration of converting phages at both sites. The observation that one of the *Bam*HI fragments hybridizing with *attB1-attB2* is the same size in all but one of the *C. diphtheriae* examined in this study, whereas the second varies in size, is still unexplained. Integration at both C7 sites has been reported (13) and would be expected to produce size variation in both fragments. It is possible that in these strains phage was preferentially integrated into one site, thus producing variation in the size of only one of the *Bam*HI fragments.

The most intriguing observation of all is the wide distribution of the *attB*-related sequence among *Corynebacterium* species and related organisms. It could be argued that the β family of phages is of ancient ancestry and has accompanied the corynebacteria as they evolved from soil and plant species to species inhabiting man and other animals. This would explain the persistence of these sites but would only do so if they were still being used by phages carrying the homologous *attP* sites. So far, β -related phages have only been identified in strains of *C. diphtheriae*, *C. ulcerans*, and *C. pseudotuberculosis* (4). Phages isolated from *C. pseudodiphtheriticum* are morphologically and antigenically distinct from β phages (15), and a *C. renale* phage, RP28, though morphologically similar to β (7) but lacking homology to β in DNA hybridization tests, did not integrate into the *attB* site present in the genome of its host (unpublished data). It is clear that much more will have to be done before the hypothesis of coevolution is adequately tested. Alternatively, it could be postulated that these sites were retained because they play some other role in the genetics, and hence evolution, of the corynebacteria, though there is no evidence in support of this hypothesis. Finally, it is of interest that, in three strains of *C. ulcerans* and one of *C. pseudotuberculosis* in which the *tox* gene is present in the absence of any other detectable β -related DNA (5), one of the *Bam*HI fragments hybridizing with the *attB1-attB2* probe also contains the *tox* gene. Whether this reflects the original relationship between *tox* and a phage integration site or a residue of some past interaction of the site with a phage has yet to be determined.

ACKNOWLEDGMENTS

We acknowledge the technical assistance of Myron Rabin and Inyong Chang in portions of this work.

This investigation was supported by Public Health Service research grant AI-10492 from the National Institute of Allergy and Infectious Diseases and by National Research Service award GM-07720 from the National Institutes of Health.

LITERATURE CITED

1. Buck, G. A., R. E. Cross, T. P. Wong, J. Loera, and N. Groman. 1985. DNA relationships among some *tox*-bearing corynebacteriophages. *Infect. Immun.* 49:679-684.
2. Buck, G. A., and N. B. Groman. 1981. Physical mapping of β -converting and γ -nonconverting corynebacteriophage genomes. *J. Bacteriol.* 148:131-142.
3. Cianciotto, N., and N. Groman. 1985. A beta-related corynebacteriophage which lacks a *tox* allele but can acquire it by recombination with converting phage. *Infect. Immun.* 49:32-35.
4. Groman, N., N. Cianciotto, M. Bjorn, and M. Rabin. 1983. Detection and expression of DNA homologous to the *tox* gene in nontoxicogenic isolates of *Corynebacterium diphtheriae*. *Infect. Immun.* 42:48-56.

5. Groman, N., J. Schiller, and J. Russell. 1984. *Corynebacterium ulcerans* and *Corynebacterium pseudotuberculosis* responses to DNA probes derived from corynephage β and *Corynebacterium diphtheriae*. Infect. Immun. 45:511-517.
6. Groman, N. B. 1984. Conversion by corynephages and its role in the natural history of diphtheria. J. Hyg. 93:405-417.
7. Hirai, K., and R. Yanagawa. 1970. Generalized transduction in *Corynebacterium renale*. J. Bacteriol. 101:1086-1087.
8. Laird, W., and N. Groman. 1976. Prophage map of converting corynebacteriophage beta. J. Virol. 19:208-219.
9. Maximescu, P. A., A. Oprisan, A. Pop, and E. Potorac. 1974. Further studies on *Corynebacterium* species capable of producing diphtheria toxin (*C. diphtheriae*, *C. ulcerans*, *C. ovis*). J. Gen. Microbiol. 82:49-56.
10. Michel, J. L., R. Rappuoli, J. R. Murphy, and A. M. Pappenheimer, Jr. 1982. Restriction endonuclease map of the nontoxic corynephage γ c and its relationship to the toxigenic corynephage β c. J. Virol. 42:510-518.
11. Rappuoli, R. 1983. Isolation and characterization of *Corynebacterium diphtheriae* nontandem double lysogens hyperproducing CRM197. Appl. Environ. Microbiol. 46:560-564.
12. Rappuoli, R., J. L. Michel, and J. R. Murphy. 1983. Restriction endonuclease map of corynebacteriophage ω_{tox} isolated from the Park-Williams no. 8 strain of *Corynebacterium diphtheriae*. J. Virol. 45:524-530.
13. Rappuoli, R., J. L. Michel, and J. R. Murphy. 1983. Integration of corynebacteriophages β_{tox} , ω_{tox} , and γ_{tox} into two attachment sites on the *Corynebacterium diphtheriae* chromosome. J. Bacteriol. 153:1202-1210.
14. Rappuoli, R., and G. Ratti. 1984. Physical map of the chromosomal region of *Corynebacterium diphtheriae* containing corynephage attachment sites *attB1* and *attB2*. J. Bacteriol. 158:325-330.
15. Saragea, A., P. Maximescu, and E. Meltzer. 1976. *Corynebacterium diphtheriae*: microbiological methods used in clinical and epidemiological investigations, p. 62-176. In T. Bergan and J. R. Norris (ed.), Methods in microbiology, vol. 13. Academic Press, Inc., New York.
16. Schiller, J., N. Groman, and M. Coyle. 1980. Plasmids in *Corynebacterium diphtheriae* and diphtheroids mediating erythromycin resistance. Antimicrob. Agents Chemother. 18: 814-821.
17. Weisberg, R. A., and A. Landy. 1983. Site-specific recombination in phage lambda, p. 211-250. In R. W. Hendrix, J. W. Roberts, F. W. Stahl, and R. A. Weisberg (ed.), Lambda λ . Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.